Using authentic assessment in satisfying CEAB’s new graduate attributes accreditation criteria

Anastassis Kozanitis, Ph.D.
Center for Faculty Development, École Polytechnique
anastassis.kozanitis@polymtl.ca

Guy M. Cloutier, ing., Doc.
Dept of Mechanical Eng., École Polytechnique
guy.cloutier@polymtl.ca

Abstract

“CEAB 2014” affects student assessment methods. There is a risk of an increased workload, as assessing for grading in a course and for the possession of attributes are “distinct matters”. Here is a rationale for assessment of learning outcomes, with an overview of methods and issues to consider when designing and using assessment of attributes in relation to CEAB’s accreditation criteria. To increase their performance, students need ‘educative assessments’, anchored in authentic tasks and with feedback usable to improve performance. At École Polytechnique de Montréal, a committee studies this issue, in both project modules and internships. The initiative aims to give students a more active role, engaging them in deep experiential learning. It faces the challenge to recommend effective and reliable assessment methods for student grading, also meeting the new requirements of CEAB with current resources. Association with other Canadian institutions is sought to continue the discussion.

1 Introduction

Canadian engineering programs are to assess and report on what is learned, by 12 graduate attributes: 1) knowledge; 2) problem analysis; 3) investigation; 4) design; 5) use of engineering tools; 6) individual and team work; 7) communication; 8) professionalism; 9) impact of engineering on society and the environment; 10) ethics and equity; 11) economics and project management; 12) life-long learning.

Accreditation bodies in various countries do their best to help schools and faculties adapt to the change, often depicted as having ‘minor implications’ on daily teaching/assessment methods and interactions. Even so, many instructors and coordinators see the change as quite challenging. Literature presents such changes as a not so trivial: a paradigmatic shift from a teaching to learning, a “revolution” focusing on combinations of knowledge, skills and attitudes students should possess, and on ways to appraise it. This shift requires innovative thinking, as well as changes in habits. The approach, inspired in part by changes initiated in 1997 by ABET [1] following the “Boeing report” of 1996 [2], has also directly or indirectly motivated initiatives to transform the way engineering schools think of preparing their students for the labour market. Programs realise that keeping with former methods will not help meet the new criteria.

A companion paper summarises an institution-wide renewal of the curricula at École Polytechnique, implemented in 2005 [3]. It affected the pedagogical design of all programs. By design, programs pay more attention to communication and teamwork, they reflect an increase in the use of active pedagogies and of experiential approaches, and they provide opportunities to develop skills effective in lifelong learning [4].

This article’s aim is to provide guidance about engineering assessment practices. It addresses the need to use authentic assessment, a relevant response to CEAB’s requirements. The first part of the paper describes authentic assessment, and explains the knots and bolts of how it can be applied. The second part presents a project module where a more authentic assessment has been implemented (to the extent of our imagination and boldness) in year one of the program.

2 Satisfying new criteria for CEAB 2014

CEAB 2014 aims to assure quality and to foster the systematic pursuit of its improvement in engineering education. Programs must demonstrate that graduates meet (distinct) threshold levels on 12 attributes by the time of graduation, and have evidence assessment results contribute to ongoing upgrading. The process must assess possession of “the 12” on some scale.

In their answers of April 14 2011 [5], CEAB clarifies (we underline, and we translate freely):

“...in theory, a student could pass a course while not meeting the threshold requirement on an attribute, or the opposite. To pass a course and to demonstrate one possesses an attribute are distinct matters”

“...attributes are assessed for the purposes of the programs, and not for students taken individually.”

We conclude the demonstration to be of a convincing character, such that CEAB and visitors would be willing to rely and act upon without hesitation, rather than “absolute certainty about every student.”
CEAB also explicitly mentioned or stressed that:

1. In the absence of CEAB’s definition of complexity, work in progress should refer to the definition provided by the International Engineering Alliance (IEA) [6].
2. Engineering schools and faculties are responsible for establishing the “threshold” and the “target” levels they will pursue for the 12 attributes.
3. Students can undertake a wide variety of (extra) curricular activities to develop attributes.
4. Appraisal contexts and situations apt at covering multiple attributes are judicious means.

Meeting the obligations can ask for adjustments in current practices. It is insufficient to replicate routines, methods, and facts in a final exam, where students largely face what Wancat calls (modified from) previously solved well-defined givens and goals ‘routine problems’ [7]. Discerning what constitutes good practice for assessment warrants ongoing efforts. Appraisal practices should be revised if we are to improve learning and preparedness for professional practice on all 12 attributes. Students need appraisals i - anchored in authentic tasks, and ii - with opportunities to improve proficiency transferable to similar tasks. Authentic assessment offers both.

3 Authentic assessment

Authentic assessment requires students to act upon open-ended situations as a professional would, by mobilizing their knowledge as well as their broad skills and attitudes. (Traditional assessment focuses on information recollection, and mobilises little if any analysing, judging and critical thinking skills.)

3.1 WHAT is authentic assessment

Authentic assessment proposes realistic situations, where mandates and goals exhibit the aspects of professional life, and where a response combines skills, knowledge, and attitudes. For Tardif [8], authentic assessment situations are contextualised and do not refer to an artificially simplified work statement. They call for judgment, so students must use knowledge and skills with caution to devise solutions and elaborate answers. The result entails creativity or innovation when compared to the standardized answers awaited in traditional assessment. Students should do more than produce answers in abstracto, replicate a sequence of steps, or blindly compute figures. Instead, they should respond to several situational, circumstanced or conditional aspects: ethics, rules and regulations, budget limits, imposed data-processing environments, and uncontrollable human behaviour. These conditions could be contiguous to the required answer, and affect the nature of the answer or of the work.

The evaluation of students then is constructive and iterative, more formally than in real life situations. Wiggins [9] lists several characteristics for authentic assessment: being realistic, requiring judgment and innovation, requiring action beyond reproduction, resembling workplace situations, and allowing consultations, feedback and progress.

Authentic assessment favours complexity as defined by the IEA (i.e. note (c) of Table 1). Solving complex problems requires proper integration of material from various courses and iterations. Feedback from the instructor improves proficiency. Assessment becomes an integral part of teaching... (no) of deeper learning.

In the semester, students will cyclically ask for feedback to clarify their vision of what excellent work is like, and increase their level of proficiency. Allowing errors and learning from them forms a key ingredient of authentic assessment. Transparent, publicly stated criteria govern feedback. Continuous feedback nurtures convergence towards a vision of proficiency. This is not unlike estimation by recursive filtering, where noisy observations finally produce an estimate closer to the true value than a single “good” measurement would. Instructors are sensors; leave the “system’s dynamic model” and the “prediction” to the student: this is learning. Frequency of feedback is key; it is not “taking the student by the hand.”

The student faces complex, (nearly) real world situations, and is assessed during and after completion of a specific mandate within. Ideally, he was informed of the mandate, goals and responsibilities entrusted to him. Instructors have clarified explicit expectations, and have set clear deadlines and constraints (material, human, financial, etc.). They assess the students “in” context, and interpret the results in the light of it.

Real world contexts, mandates and repeated assessments hold key relations to the 12 attributes. These essentials can only suffer minor adjustments because of program constraints on resources. Hands-on projects offer a natural setting for authentic assessment in engineering education. Nevertheless, authentic assessment requires well organized planning to help design adequate assessment instruments and effective feedback mechanisms, likely to foster learning and success. Hence, implementing authentic assessment requires formulating clear achievement targets, choosing real-world tasks, specifying a complex problem, and giving clear set of instructions.

3.2 WHY use authentic assessment of attributes

CEAB explicitly admits multiple attribute assessments can take place on a single appropriate occasion. Given the strong interrelation between some attributes within a complexity – as for Problem analysis and Investigation, or for Design and Use of engineering tools –, there may be no choice but to assess a number
of attributes on a single occasion. Only (nearly) real world situations are likely to offer a context appropriate to the expression of numerous attributes simultaneously and interactively.

CEAB asks each of the 12 attributes be assessed distinctly. By the confounding and compounding effects of attributes within any single assessment instrument, distinct assessment of the 12 attributes is only made possible by repeated observations through numerous instruments, and by a posteriori inductive reasoning. Here again, only the multiplicity of the repeated observations of authentic assessment are likely to provide more than a projection in a subspace where numerous attributes are confounded.

### 3.3 HOW to use authentic assessment of attributes

We propose a four-step process for designing effective authentic assessment of attributes and skills: 1) Organizing and planning the process; 2) Choosing or conceiving an assessment instrument; 3) Foreseeing feedback mechanisms; and (trying to) 5) Continuously improve the process.

#### 3.3.1 Organizing and planning the process

Considering the factors of the assessment situation and analyzing their influence on assessment is done by answering a few questions: What are the attributes or the skills that are needed to be assessed? What is the expected level of achievement for each attribute? When in the curriculum and during what (credited or not) activity to assess? Who will assess (and have the competencies to do so)? Which students and how many students will be assigned to each assessor? What dimensions of complexity should the situation present and to what extent? Which specific professional actions/descriptions/behaviours are expected?

The answers define the profile of the assessment exercise. For instance, the level of proficiency will differ both between attributes and longitudinally along the curriculum. So will the expression of complexity. The number of students to assess will affect the extent of the assignment and the size of the teams. The “most adapted combination” of assessment instruments results from an adequate compromise.

#### 3.3.2 Choosing or designing an instrument

Assessment instruments fit within a taxonomy of authenticity. Not so authentic instruments include multiple choice exams, short answer exams, cookbook type lab sessions or exercises, drill, and theoretical proofs. In *educating for the practice of a profession*, these may suit some of the knowledge base upstream. They should no longer be used for professional skills and (quasi)competencies. Team projects, work simulations, case studies, long answer exams within a professional setting, artistic creations, mock trials, and clinical evaluations are more authentic instruments. Their features are better suited for the assessment of attributes; yet do not produce grades so easily. They ask for deeper involvement from the part of the assessor: a master student parachuted at the end of the semester will have a tough uncomfortable time writing down marks, whatever the rubric used.

With Wankat [7], not-so-well-defined problems can be sorted by the vector of “degree-defined” (givens & goals), “previous exposure” (replicate, modified from, never seen), “types of unknown” (design, cause and cure, understanding, discovery), and “problem solving taxonomy” (routine, diagnosis, strategy, interpretation, generation). Extrapolating from problem to situation, authentic assessment situations need little more.

The more the situation is undefined, never seen, open-ended, strategic or interpretative, the more students will need guidance to know what is expected from them: “crystal ball gazing” is not an attribute for CEAB. Mandates, goals and deliverables are provided. Students’ organisation and deliverables is what is being assessed. At lower levels of proficiency, steps to follow and details of the schedule to respect could be supplied. At higher levels or when assessing teamwork, the breakdown in steps, the sharing of tasks and the common scheduling may be the deliverables.

Assessing complex situations is a delicate exercise because of the possible bias due to the subjectivity of human judgment. It only becomes a problem when judgment is based on fuzzy or undisclosed criteria. That is where rubrics explicitly listing the criteria come in handy: distinguishing between levels of proficiency, and allowing complementary self- or peer-assessments (whether or not these affect grading). Characteristics of good rubrics are presented in [10].

#### 3.3.3 Foreseeing feedback mechanisms

Feedback about work in progress is an important part of the global assessment process, mostly for students to develop insight about their learning. As in control theory, closing the loop frequently and in a timely fashion allows students to reinforce expected attitudes and skills, and adjust for improper ones. Stressing facts, repercussions and openness for improvement should be the perspective taken by instructors when providing feedback: “closing the loop” is not “setting grades”. Feedback remarks are to be detailed, precise, and constructive, and refer to the instructions given, to the intended outcomes and to the relevant attributes.

Huba and Freed [10] distinguish between intermediate and final feedback. Rubrics are usable in both, provided “marks” are not indiscriminately associated to rubrics in the minds of students: rubrics
have to help students foresee standards by which they will be evaluated, and characteristics they will be graded on. Feedback discussions should be scheduled in a timely manner, including after final assessment: reflecting of assessment is part of the learning process. Either oral or in writing, feedback looks at sharing information and exploring alternatives, not advising or imposing solutions: using feedback to make headway is part of the skills/attitudes to nurture in the students.

Questioning students is a powerful tool to help students develop an understanding of what they have learned, and what they still need to work on. Open-ended questions help develop their critical thinking, as well as their metacognitive abilities. Metacognition is a most important ingredient of the life-long learning attribute. Open-ended questions could be: Why choose \( x \) and not \( y \)? How do you know this is the right way? What was the hardest thing to overcome in order to ...? What are the implications of your conclusions to \( x \)? How can you improve the \( x \) process, using the \( y \) model?

3.3.4 Continuously improving the process

Assessing the assessment process and looking for areas of improvement closes the loop for the instructor: validity, reliability and robustness are genuine sources of questioning. The following pen-ended questions can help with the improvement phase: Do the selected instruments allow adequate assessment of the targeted attributes? Are the tasks to be realized clearly presented? Is the rubric clear, univocal, and complete? Does the assessment process as a whole allow feedback and the continuous improvement of work by the students? Do students receive sufficient support? Other methods encompass i. to analyse students’ results and try to identify replications in difficulties that could indicate eventual weaknesses in the global process; ii. to meet a few students and discuss potential improvements, which will rapidly point to very concrete and operational aspects drawn from the students’ experience.

4 A first year project module with a more authentic assessment

This project module is a first opportunity for students to confront a priori beliefs and expectations about design and teamwork, with “reality” (inasmuch as a first-year module can use “reality” as a learning/assessment context).

4.1 Global description

This second semester project module of 3 credit-points (135 hours total per student) groups students in teams of five (674 person-hours, rough, from which to subtract large overlaps given the ‘learning’ context). Some of the basic courses it follows are in Statics, in Technical drawing and in Industrial psychology. Amongst others, it accompanies courses in Strength of materials, in Mechanical dimensioning and in Computer programming. In parallel to ‘courses’, coaching on formal oral and written communication is maintained as a longitudinal process up to year three.

The module is composed of (per team, nine teams):

- an introductory Design-Build-Test criteria-based common experience,
- three case studies chosen along the design phases of a product life-cycle
- and three formal reports,
- three criteria and general-rubric guided peer-assessments (one per case study),
- three team-contract reviews and negotiations,
- three quick mechanical dissection exercises introducing the distinctive aspects of the cases,
- three interactions with panels of P.Eng.,
- six literature searches (theory, processes, standards,...) related to the case studies
- and six one-page summary exercises,
- three ‘problem mapping’ and work scheduling, with diminishing guidance,
- three formal oral presentation per team, either on the case study itself or on complementary work,
- 3 x 9 criteria-based and detailed-rubric guided peer-assessments of oral presentations,
- six appraisals of logbooks over the semester,
- and one final evaluation of logbooks per student,
- 10 coordination meetings in all, each with its secretary, agenda, and meeting report,
- 10 workgroups within contact-hours, with the instructors observing, questioning and answering,
- 10 workgroups outside of contact-hours,
- two visits of teamwork specialists,
- a closing competition about team organisation on a hands-on challenge ($100 per student for the winning team).

All teams are subjected to random selection for defence of work in progress, with feedback from the class and from the instructors.

4.2 Scenario for appraisal

The module aims at an authentic contribution to Problem analysis, Investigation, Individual and teamwork, and Life-long learning. Although the context is design, our perception is the fragmentation into cases takes much of the ‘design process responsibility’ away from the students for a serious contribution to the Design attribute. From the list in section 4.1, Communication (written and oral, formal and informal) is obviously omnipresent.
Test benches for a car scissors-jack and the car jack itself make the ‘design context’. Specifications, together with car brands and models, are made to vary. As students take team roles in rotation, meeting with the team leaders both mimics a (soft) chain of command, and provides initial status to the new temporary leader. Formally addressing the team is done through the secretary (of the case study). Roles cannot be held twice by the same student.

The CEAB attributes present were fragmented into constituents with the CDIO Syllabus [11], aiming at level 2: “Participates and contributes to...”, and level 3: “Executes with signs of judgement...” for most but ‘problem identification’ where level 1: “Has been exposed to...” is deemed satisfactory. These are ‘modified CDIO proficiency levels’, a matter discussed in the companion paper (see [3]). The most relevant aspects of this fragmentation are presented in Table 1, per CEAB attribute. The table also displays transversal skills beyond those specified by CEAB, and the aspects of complexity ‘experienced with’ as defined by the IEA. Complexity is to be understood ‘as the student stands in the curriculum’.

*Life-long learning* warrants additional explanations.

As the semester progresses, students fall short from the knowledge required to address *buckling* of the U-channels of the scissors jack. By forcing this analysis, students are put in a situation for which their knowledge is insufficient.

To us, recognising the limits of one’s knowledge, finding the knowledge requirements of a need, finding sources of information, and taking the responsibility to learn and to apply the new knowledge are all essentials of life-long learning. To reflect on what needed to be learned, how it was learned, and on aspects of surface or deep learning completes the process.

The presentation of the numerous rubrics and tools used is beyond the scope of the present paper.

5 Conclusion

Authentic assessment of attributes sets itself apart from traditional practices, which are too often satisfied with the repetition of awaited good answers and routines. Authentic assessment calls for the design of instruments capable of evaluating students’ attributes in realistic and complex situations. These simulate real life professional mandates. Producing the assessment realistic and complex situations. These simulate real instruments capable of evaluating students’ attributes in

A first-year project module was progressively improved since a major renewal in 2005. It exemplified the structure of a favourable context. It showed how a breakdown of CEAB attributes helps to manage them with the granularity level found in a course. Finally, it displayed the frequent and intensive nature of authentic assessments and feedback, likely to ‘make a difference’ in the learning process of the students. A rather strong departure from traditional assessment.

6 References


Table 1. Attributes of first year integration project module with relevant aspects of transversal skills and of complexity, prerequisites of knowledge, with experienced and observed knowhow. Vector of expression includes direct observations, logbooks, meeting notes, oral presentations and defense, written report, videos, computer codes.

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Aspects</th>
<th>Components perceived and retained, directly or indirectly observed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Problem analysis</td>
<td>Transversal</td>
<td>Critical thinking (evidence, arguments, theories and facts, hypotheses and conclusions); Communication (formal written and oral)</td>
</tr>
<tr>
<td>Complexity</td>
<td></td>
<td>Issue infrequently encountered; Many component parts with different optima (never seen before)</td>
</tr>
<tr>
<td>Knowhow</td>
<td>Statics; Strength of materials; Procedural programming; Drawing</td>
<td></td>
</tr>
<tr>
<td></td>
<td>To formulate; To model (conceptual, qualitative, quantitative, numerical); To estimate (orders of magnitude, consistency); To analyse in the presence of uncertainty (incomplete and ambiguous information, sensitivity to chosen/given parameters); To solve and recommend (design specifications).</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Transversal</td>
<td>Creative thinking (association); Resourcefulness; Will to deliver</td>
</tr>
<tr>
<td>Complexity</td>
<td></td>
<td>Not covered by protocols, procedures, and instruments (as the students stand in the curriculum)</td>
</tr>
<tr>
<td>Knowhow</td>
<td>Phenomenological (Hooke’s law, Statics); Methodological (nil); Technical (safety only)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>To formulate (critical questions); To inquire by experiment (strategy, precautions, construction, procedures, measurements, data vs. available models; all in the presence of an hyperstatic system); To test and defend results obtained by induction (informal defense)</td>
<td></td>
</tr>
<tr>
<td>Investigation</td>
<td>Transversal</td>
<td>Temperaments (MBTI); Attitude (awareness of self, of others and of interactions, reflection, loyalty and dependability, curiosity, openness, listening and dialog); Involvement (mutual trust, feedback and negotiation, compromise and conflict resolution); Personal communication (verbal, non verbal)</td>
</tr>
<tr>
<td>Complexity</td>
<td>Requirements on team resources perceived as conflicting (as students stand in the curriculum)</td>
<td></td>
</tr>
<tr>
<td>Knowhow</td>
<td>Temperaments; Experiential learning; Basic industrial psychology</td>
<td></td>
</tr>
<tr>
<td></td>
<td>To decide and act in the face of uncertainty (opportunities, potential in spite of risks, initiating tasks and acting on ideas, delivery of results); To persevere, be resourceful, flexible, responsible (self-confidence, enthusiasm, importance of hard work, intensity, attention to detail, ability to work independently); To overcome obstacles (extent of one’s responsibility, willingness to complement the work of others, to consider others’ viewpoints in own work, positive response to criticism); To manage time and resources (recognition of sub-tasks, prioritisation, allocation of time and resources, importance and/or urgency of tasks, efficient execution)</td>
<td></td>
</tr>
<tr>
<td>Individual and team work</td>
<td>Transversal</td>
<td>Individual knowhow</td>
</tr>
<tr>
<td>Complexity</td>
<td>Temperaments; Experiential learning; Vulgarised/elementary neuro-physiology (attention, concentration, short and long term memory, deep learning)</td>
<td></td>
</tr>
<tr>
<td>Knowhow</td>
<td>To form effective teams (team goals, processes and roles, needs of members, strengths and weaknesses of members and of the team, confidentiality, accountability, initiative); To operate as a team (agenda, planning, meeting, communicating, positive and effective feedback, project planning, scheduling and execution, decision making, resolution of conflicts); To grow and evolve (peer assessment and self-assessment, cohesion, mutual care, synergy); To make headway (objectives, process management, facilitation styles, approaches to motivation, representing the team, mentoring and counseling)</td>
<td></td>
</tr>
<tr>
<td>Team knowhow</td>
<td>To identify requirements, recognise own limits, and be capable of knowledge integration (indicators about knowing, knowing to do and knowing to be, mechanisms of deep learning); To learn by self and through others (curiosity, motivation for self-education, being mentored, master own learning style and compensate for its drawbacks)</td>
<td></td>
</tr>
<tr>
<td>Life-long learning</td>
<td>Transversal</td>
<td>Will and desire to maintain/improve significance, effectiveness, efficiency and other talents/aptitudes; Responsibility for self-improvement, and to overcome important weaknesses</td>
</tr>
<tr>
<td>Complexity</td>
<td>(N/A)</td>
<td></td>
</tr>
<tr>
<td>Knowhow</td>
<td>Experiential learning; Basic industrial psychology</td>
<td></td>
</tr>
<tr>
<td></td>
<td>To identify requirements, recognise own limits, and be capable of knowledge integration (indicators about knowing, knowing to do and knowing to be, mechanisms of deep learning); To learn by self and through others (curiosity, motivation for self-education, being mentored, master own learning style and compensate for its drawbacks)</td>
<td></td>
</tr>
</tbody>
</table>

(a) **Critical thinking**: Observables: issue; evidence; arguments; perspectives, theories, and facts; fallacies; hypotheses and conclusions. Metrics: relevancy/appropriateness; logic; completeness/coverage. Vector of expression: written proofs and demonstrations; oral defense; reports; case studies; ...

(b) **Creative thinking** – Observables: associations between seemingly independent aspects; conceptualisations and abstractions; synthesis and generalization. Metrics: frequency; power; breadth; depth. Vector of expression: written proofs and demonstrations; oral defense; reports; case studies; ...

(c) **Complexity**: conflicting requirements; abstract thinking, and originality of model; fundamentals-based, first principles analytical approach; issues infrequently encountered; not covered by standards and codes; involve diverse groups of stakeholders; with significant consequences in a range of contexts; with many component parts or sub-problems.

† To modulate along the curriculum (from one to three simultaneous dimensions of complexity – one in the first and the second year), and to adjust or to put in the perspective of expectations ‘as the student stands in the curriculum’ (e.g. “infrequently encountered issues” up to the second year of the curriculum).